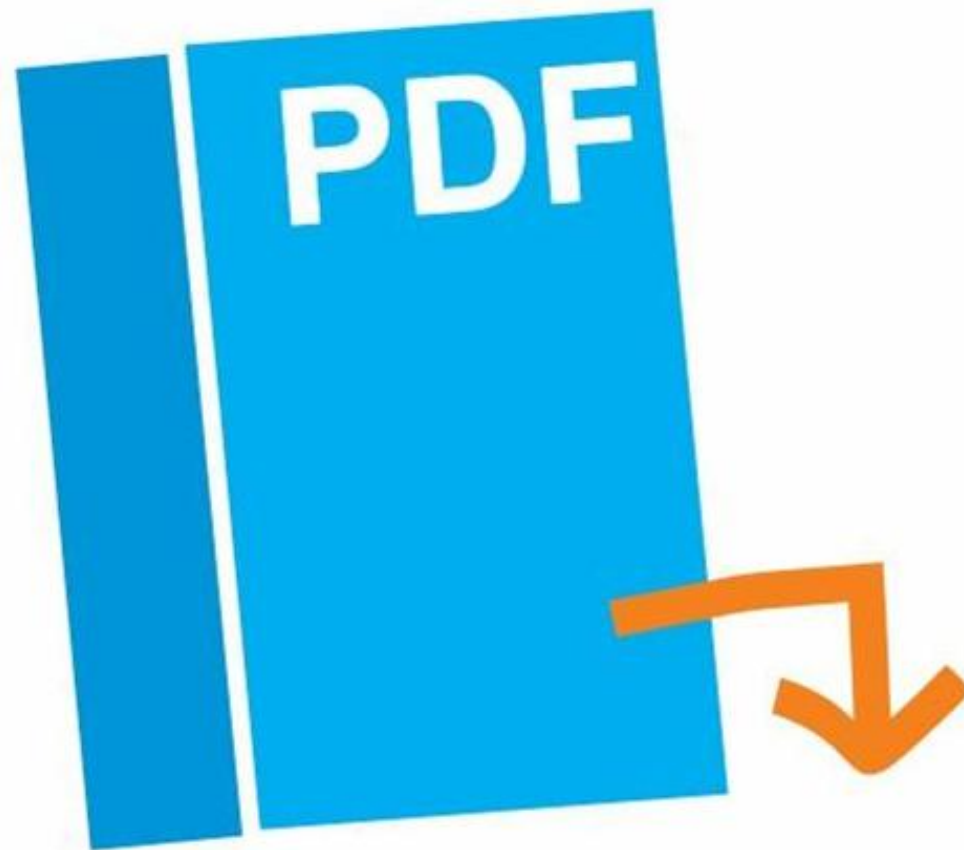


## Полное название:

«Разработка веб-приложения для каталогизации и просмотра файлов в формате PDF с автоматическим распознаванием их содержимого»

## Краткое название:

«ПДФокус»



Выполнили: Каткова А. А.,  
Тимерзянов Т. Н. | Научный  
руководитель: Ступников А.А.

## Цель

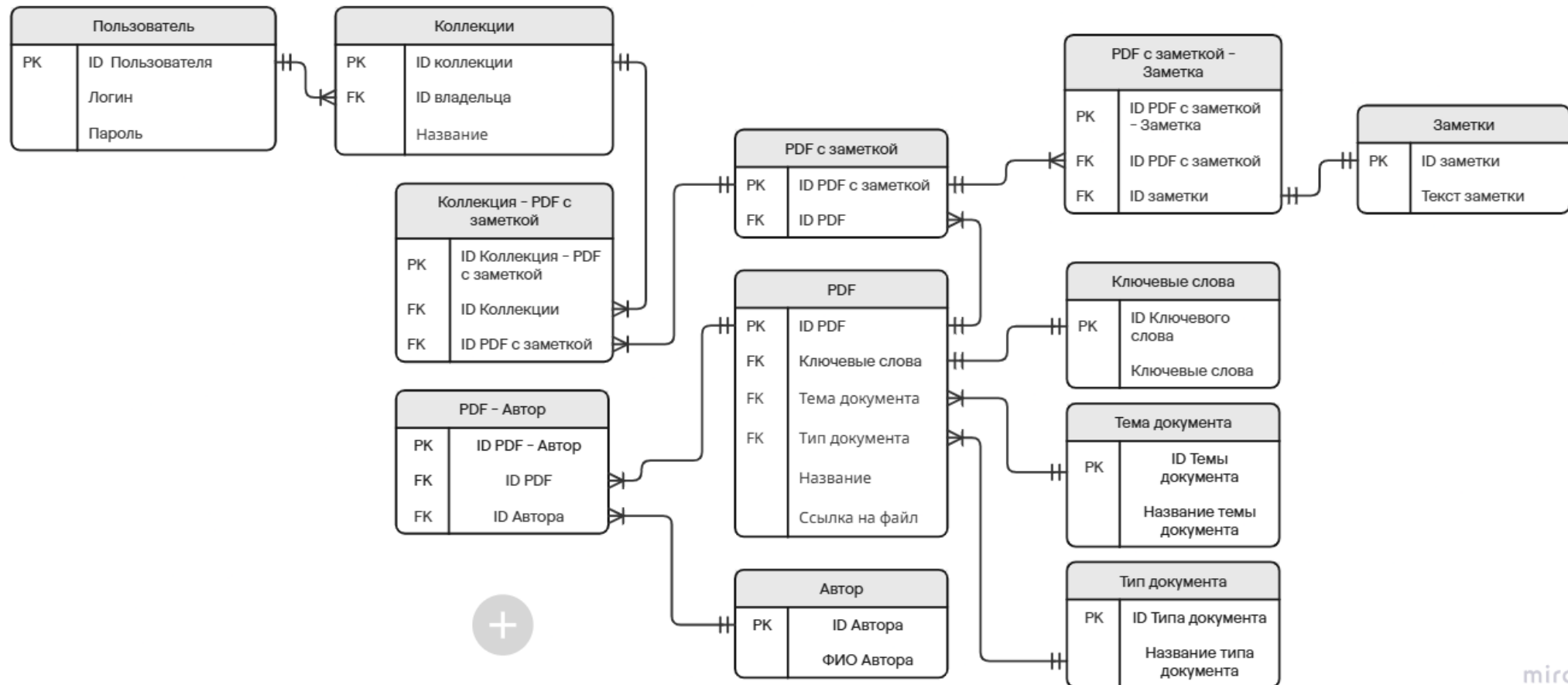
- Разработка веб-приложения, позволяющего помещать в личное облачное хранилище важные PDF-файлы, просматривать их, автоматически каталогизировать через алгоритмы МО и проводить полнотекстовый поиск по содержимому с помощью OCR

# Задачи

- ❑ Исследовать технологии для работы с PDF-файлами
- ❑ Проанализировать и выбрать алгоритмы для автоматического распознавания текста (OCR)
- ❑ Разработать архитектуру веб-приложения
- ❑ Реализовать модуль для загрузки, хранения и каталогизации PDF-файлов.
- ❑ Интегрировать компонент для распознавания текста с возможностью получения структурированной информации из загруженных файлов.
- ❑ Провести тестирование пользовательского интерфейса и функциональности на разных устройствах.

# Информационная составляющая

## PostgreSQL для записи мета-данных о PDF-файлах



# Функциональная составляющая

- ☐ Загрузка и просмотр PDF-файлов
- ☐ Поиск по каталогу PDF-файлов
- ☐ Распознавание текста из загруженных PDF-файлов
- ☐ Управление личным кабинетом
- ☐ Возможность добавления заметок и аннотаций

# Render.com — ХОСТИНГ

My Workspace

Dashboard

PDFocus

Events

Settings

MONITOR

Logs

Metrics

MANAGE

Environment

Shell

Scaling

Previews

Disks

Jobs

Changelog

Invite a friend

Contact support

Render Status

PDFocus

WEB SERVICE

PDFocus

Python 3

Free

Upgrade your instance →

Connect

Manual Deploy

CodeOfArtemis / PDFocus

Artem\_branch

https://pdfocus.onrender.com

Your free instance will spin down with inactivity, which can delay requests by 50 seconds or more.

Upgrade now

June 9, 2025 at 9:37 PM

Live

2c92a89

Add media files serving in production

All logs

Search

Live tail

GMT+5

AppleWebKit/537.36 (KHTML, like Gecko) Chrome/137.0.0.0 Safari/537.36"

Jun 9 10:02:24 PM 127.0.0.1 - - [09/Jun/2025:20:02:24 +0300] "POST /auth/ HTTP/1.1" 302 0 "https://pdfocus.onrender.com/auth/" "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/137.0.0.0 Safari/537.36"

Jun 9 10:02:25 PM 127.0.0.1 - - [09/Jun/2025:20:02:25 +0300] "GET /main/ HTTP/1.1" 200 3892 "https://pdfocus.onrender.com/auth/" "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/137.0.0.0 Safari/537.36"

Jun 9 10:04:37 PM 127.0.0.1 - - [09/Jun/2025:20:04:37 +0300] "GET /catalog/ HTTP/1.1" 200 6117 "https://pdfocus.onrender.com/main/" "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/137.0.0.0 Safari/537.36"

Jun 9 10:04:38 PM 127.0.0.1 - - [09/Jun/2025:20:04:38 +0300] "GET /static/css/styles.css HTTP/1.1" 304 0 "https://pdfocus.onrender.com/catalog/" "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/137.0.0.0 Safari/537.36"

Jun 9 10:04:38 PM 127.0.0.1 - - [09/Jun/2025:20:04:38 +0300] "GET /static/css/auth.css HTTP/1.1" 304 0 "https://pdfocus.onrender.com/catalog/" "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/137.0.0.0 Safari/537.36"

Jun 9 10:04:38 PM 127.0.0.1 - - [09/Jun/2025:20:04:38 +0300] "GET /static/js/app.js HTTP/1.1" 304 0 "https://pdfocus.onrender.com/catalog/" "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/137.0.0.0 Safari/537.36"

Jun 9 10:04:38 PM 127.0.0.1 - - [09/Jun/2025:20:04:38 +0300] "GET /static/js/simple-pdf-viewer.js HTTP/1.1" 304 0 "https://pdfocus.onrender.com/catalog/" "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/137.0.0.0 Safari/537.36"

Jun 9 10:20:38 PM [2025-06-09 20:20:38 +0300] [84] [INFO] Handling signal: term

Jun 9 10:20:38 PM [2025-06-09 20:20:38 +0300] [90] [INFO] Worker exiting (pid: 90)

Jun 9 10:20:40 PM [2025-06-09 20:20:40 +0300] [84] [INFO] Shutting down: Master

Need better ways to work with logs? Try the Render CLI or set up a log stream integration →

Добро пожаловать!

Главная

Каталог PDF

### Организуем MLP для задачи классификации так:

**1. Число задаваемых классов можно брать любое.**

*Это достигается организацией в выходном слое отдельного нейрона для каждого варианта.*

**2. Сохраним наличие только 1 скрытого слоя.**

**3. Число нейронов каждого слоя может быть любым.**

*Организуем 3 массива нейронов.*

**4. Реализуется более быстрый алгоритм обратного распространения ошибок.**

*Организуем возможность использования мини-батчей.*

**5. Реализуется контроль переобучения на каждой эпохе.**

*Для этого обучающая выборка разбивается на 2 части: обучающую (большую) и контрольную (малую).*

*По обучающей части выборки выполняется классическое обучение. Качество обучения проверяется по обеим выборкам.*

Добро пожаловать!

ПДФокус

Поиск по названию, автору, теме, клн

Личный кабинет

Выйти

Главная

Каталог PDF



anzhelika2

Пользователь с 09.06.2025

Файлов: 3

Заметок: 2

## Последние файлы

МТГО\_07\_Рекуррентные  
нейронные сети.pdf

anzhelika2 • Открыт: 09.06.2025 20:49

Ключевые слова: состояние, цепи,  
последовательности, уч...

Открыть

Удалить

МТГО\_06\_Сверточные сети\_.pdf

anzhelika2 • Открыт: 09.06.2025 20:48

Ключевые слова: свёрточной,  
нейронные, пример, фильтра...

Открыть

Удалить

МТГО\_05\_Распознавание  
цифр.pdf

anzhelika2 • Открыт: 09.06.2025 20:47

Ключевые слова: скрытого, выходного,  
слоя, число, весов...

Открыть

Удалить

## Последние заметки

Другое представление RNN МТГО\_07\_Рекуррентные нейронные сети.pdf Страница 15

Схема RNN МТГО\_07\_Рекуррентные нейронные сети.pdf Страница 9



Добро пожаловать!

ПДФокус

Поиск по названию, автору, теме, клн

Личный кабинет

Выйти

Главная

Каталог PDF

## Каталог PDF-файлов



МТГО\_07\_Рекуррентные  
нейронные сети.pdf

anzhelika2

09.06.2025

**Ключевые слова:** состояние, цепи,  
последовательности, учитывать,  
нейронные, ...

Открыть



МТГО\_06\_Сверточные сети\_.pdf

anzhelika2

09.06.2025

**Ключевые слова:** свёрточной,  
нейронные, пример, фильтра,  
повходному, наиболе...

Открыть



МТГО\_05\_Распознавание  
цифр.pdf

anzhelika2

09.06.2025

**Ключевые слова:** скрытого,  
выходного, слоя, число, весов,  
обучающей, обоих, ...

Открыть

[Сохранить изменения](#)[Скачать PDF](#)[Опубликовать](#)

### Просмотр PDF:

Редактор: Anne Schneider

Перевод: Landor

Корректор: Н. Журавлёва

© 2015 Stephen Dewey

Все права защищены

*Первое издание*

**Отказ от ответственности:** Игра *Десять свечей* предполагает использование свечей. Использование свечей и открытого огня опасно и может причинить травмы, порчу имущества и даже стать причиной смерти. Пожалуйста, соблюдайте максимальную осторожность, если будете использовать свечи. Ни при каких условиях издательство Cavalry Games и его сотрудники не несут ответственности за полученные травмы и порчу имущества (включая случайный и косвенный ущерб, персональные травмы и смерть по неосторожности, но не ограничиваясь ими), произошедшие из-за прямого или косвенного использования этих материалов, советов, руководств, рекомендаций и правил этой игры.

Пожалуйста, играйте ответственно. Огонь — это опасно.

### Текст страницы 2:

Десять свечей

Автор: Stephen Dewey

Издательство: Cavalry Games

[www.cavalrygames.com](http://www.cavalrygames.com)

Иллюстрации: J.C. Park (передняя обложка и содержание)

Stefan Harris (страницы 7, 30 и 43)

Scout Wilkinson (страница 9 и задняя обложка)

Верстка: Stephen Dewey

Редактор: Anne Schneider

Перевод: Landor

Корректор: Н. Журавлёва

© 2015 Stephen Dewey

Все права защищены

*Первое издание*

**Отказ от ответственности:** Игра *Десять свечей* предполагает использование свечей. Использование свечей и открытого огня опасно и может причинить травмы, порчу имущества и даже стать причиной смерти. Пожалуйста, соблюдайте максимальную осторожность, если будете использовать свечи. Ни при каких условиях издательство Cavalry Games и его сотрудники не несут ответственности за

# Веб-приложение

<https://pdfocus.onrender.com>

Вход

Регистрация

Имя пользователя:

Anzhelika

Пароль:

.....

Войти

